學號:R06942098 系級:電信碩一 姓名:曾柏偉

1. (1%)請比較有無 normalize(rating)的差別。並說明如何 normalize. (collaborator:)

我第一天實作時即採用了 normalize,對 rating 減去他的 mean 並且除以它的 standard Deviation,我的效果反而是變比較差的,以下為數據:

同樣都是 latent dimension: 256 且有加上 bias 的 model

loss 從

0.85869 上升至 0.86204

2. (1%)比較不同的 latent dimension 的結果。

(collaborator:)

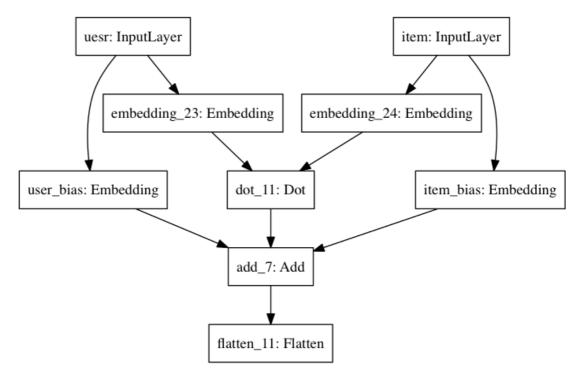
此題我同時加上了 bias,並且 latent dimension 分別為 128,256,512

Dimension	128	256	512
Loss	0.86168	0.85869	0.85949

故我都取 latent dimension 256 在 kaggle 上傳。

3. (1%)比較有無 bias 的結果。

(collaborator:)



在相同 latent dimension 2562 的情况下:

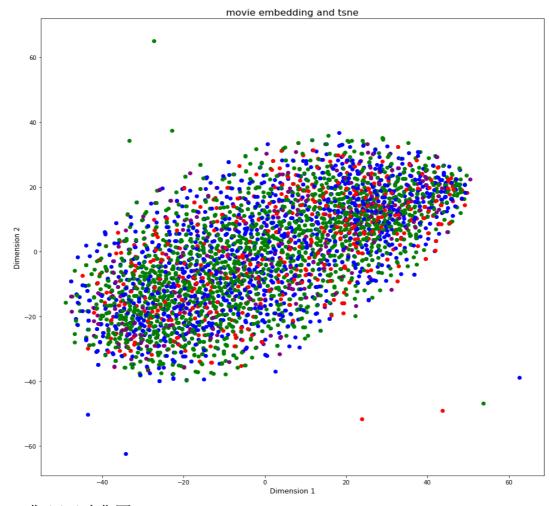
	有 bias	無 bias
Dimension: 256	0.85869	0.86290

4. (1%)請試著用 DNN 來解決這個問題,並且說明實做的方法(方法不限)。 並比較 MF 和 NN 的結果,討論結果的差異。

(collaborator:)

我實作的 DNN 一直沒辦法和助教說的一樣過 strong baseline,loss 一直卡在 0.86204 那邊 ,我採用和 sample code 相同 fully connected 的參數 (150,50,1),但是並沒有得到較好的 loss。

5. (1%)請試著將 movie 的 embedding 用 tsne 降維後,將 movie category 當



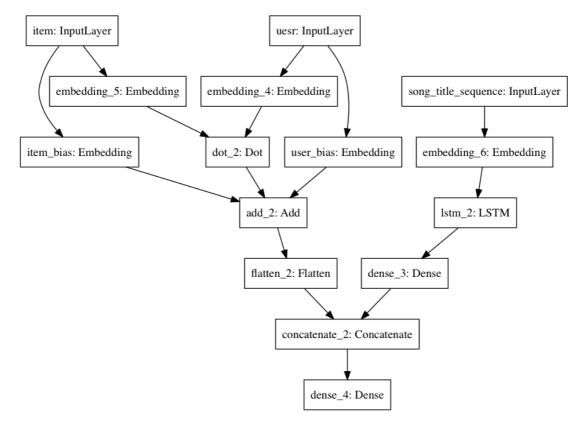
作 label 來作圖。

(collaborator:)

我依照投影片裡面,"thriller, horror, crime"為紅色點 "Drama, Musical"為藍色點 "Adventure, Animation, Children's "為紫色點

6. (BONUS)(1%)試著使用除了 rating 以外的 feature, 並說明你的作法和結果,結果好壞不會影響評分。

(collaborator:)



以上 song title 為我將 movie 的名稱對應每一個 user rating 的 data 加入進去,發現 loss 並沒有下降反而更糟了 (0.88609),下降了 0.03 左右,之後就沒有 把 user 的 feature 加進去了。